DOCUMENT RESUME

ED 423 271 TM 029 076

AUTHOR Byrne, Barbara M.
TITLE Testing for Causal Predominance between Academic Self-concept and Academic Achievement: A Developmental Perspective.
PUB DATE 1998-04-00
NOTE 39p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Diego, CA, April 13-17, 1998).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; *Causal Models; Child Development; *Elementary School Students; Elementary Secondary Education; English; *Etiology; Grade 11; Grade 3; Grade 7; *High School Students; Mathematics; *Self Concept
IDENTIFIERS Academic Self Concept

ABSTRACT

This study is an attempt to determine if it is possible to establish causal direction between academic self-concept and academic achievement, focusing on whether academic self-concept and the subject-matter self-concepts of English and mathematics serve to generate academic, English, and mathematics achievements, respectively. This study of causal predominance was tested at grades 3, 7, and 11. Parental consent was received for 252 grade 3 students, 290 grade 7 students, and 335 grade 11 students. Students at each grade level completed measures of self-concept and data relevant to academic achievement and academic self-concept were collected. Of the nine sets of parameters tested for causal predominance (three at each grade level), direction of cause was established for six. However, inconclusive findings resulted for English self-concept and English achievement relations for children in grade 3, and for academic self-concept and achievement relations and mathematics self-concept and mathematics achievement relations in grade 7. For both grades 3 and 11, tests of causal predominance for academic self-concept/achievement relations and mathematics self-concept/achievement relations demonstrated a clear flow of causality from achievement to self-concept. However, for English, a reverse pattern was demonstrated (from English self-concept to English achievement). Possible explanations are explored for results that suggest that direction of cause may be a function of subject area rather than, or in addition to, age. (Contains 4 tables, 3 figures, and 50 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 423 271

Testing for Causal Predominance Between Academic Self-concept and Academic Achievement: A Developmental Perspective

Barbara M. Byrne

University of Ottawa

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Barbara Byrne

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, April 1998

TM029076

Testing for Causal Predominance Between Academic Self-concept and Academic Achievement: A Developmental Perspective

That the link between self-concept (SC) and academic performance has provided a constant source of intrigue for researchers is evidenced by the plethora of studies devoted to the topic over the past 30 years or so (see Hattie, 1992). Questions raised by this research have been: (a) What is the relation between academic SC and academic achievement (AA)?; (b) Is there evidence of causal predominance between SC and AA?; and (c) what are the social referents that underlie the formation of SC as it relates to AA? We now review findings related to each of these issues.

Causal Predominance Between Self-concept and Academic Achievement

Of all the issues addressed concerning SC and AA relations, the one that has been the most perplexing and illusory has been the question of whether academic SC causes AA, or whether instead, AA causes academic SC. In reviewing such claims of causal predominance, Byrne (1984) found that for every study that argued for the impact of academic SC on AA, there was a comparable one that claimed the reverse to be true. Of the 23 studies reviewed, 11 argued for causal flow from SC to AA, 11 for an AA to SC flow, and one was unable to determine direction. Furthermore, in her 1984 review, Byrne noted that three other review papers were also split in their conclusions regarding this causal issue. Following a review of published studies and 18 doctoral dissertations concerned with the impact of intervention programs on the SC and AA of school children, Scheirer & Kraut (1979) reported no evidence of causal connection between the two constructs; these same conclusions were drawn earlier by Rubin, Dorle, and Sandidge (1977). In sharp contrast, however, the West et al. (1980 p. 194) review concluded that "findings are now sufficient to indicate that school achievement is 'causally predominant' over SC of academic ability". Two later studies have similarly produced mixed findings. Whereas Pottebaum, Keith, and Ehly (1986) found neither SC nor AA to be causally predominant, Skaalvik and Hagtvet (1990) reported findings of causal flow from

AA to SC based on one sample, and from SC to AA based on the other. Taken together, this melange of claims provides clear evidence that causal predominance remains yet an unresolved issue.

One major limitation of the above studies, however, is that with the exception of four, all were cross-sectional studies and therefore failed to meet the necessary criterion of temporal precedence in the determination of cause (Cook & Campbell, 1979; Kenny, 1979). This criterion demands that to assess, for example, whether SC influences AA, SC must precede AA in time. In order to meet this criterion, then, a longitudinal research design is required. One type of longitudinal design that has been used to untangle the SC/AA predominance issue is the cross-lagged panel correlation (CLPC) design (see Cook & Campbell, 1979); a literature review reveals four such studies, two of which were noted in the earlier Byrne (1984) review. Of these four studies, two reported the causal predominance of academic SC over AA for elementary (Chapman, Cullen, Boersma, & Maguire, 1981) and secondary students (Moyer, 1980), respectively; one reported the causal predominance of AA over SC for secondary students (Calsyn & Kenny, 1977), and one (Watkins & Astilla, 1987) was unable to establish any causal direction for secondary students. Here again, we find no consistency in reported findings. However, in explaining this discrepancy, we could attribute it to the known limitations associated the CLPC design (For an extended review see e.g., Byrne, 1984; Rogosa, 1980).

A statistically more sophisticated approach to the analysis of longitudinal data than cross-lagged panel analysis is the use of structural equation modeling (SEM) procedures. We now review studies that tested for causal direction between SC and AA over two or more time points using the SEM approach.

Based on two-wave data for seventh and eighth grade students, Shavelson and Bolus (1982) reported causal direction flowing from academic SC to AA. Although this academic SC/AA flow replicated for each of three subject-matter

SCs (english SC, math SC, science SC), the authors cautioned that the size (N=99) and nature of their sample (time lag of 4 months; sample from one school) warranted only tentative generalization to the population. Despite a much larger sample of high school students (N=929), data from two schools, and a longer time span (7 months), however, Byrne (1986) was also unable to demonstrate evidence of causal predominance. In a third study that spanned three time points, (Newman, 1984) tested for causal predominance between mathematics AA and math SC in second, fifth, and tenth grades. His results showed that, although AA in second grade led to changes in math SC in grade five, this causal effect diminished between grades five and ten. Following a critique and reanalysis of these data, however, Marsh (1988) argued for the flow from math SC to math achievement. In a review of these three SEM studies, Marsh (1990c) brought to light an interesting fact regarding the academic achievement indicators which may have led to the causal results reported. He noted that, whereas Shavelson and Bolus (1982) inferred AA from grades and concluded causal direction to flow from academic SC to AA, Newman (1984) inferred AA from standardized test scores and argued for causal flow from AA to academic SC. Finally, Byrne (1986) inferred AA from two measures (grades, standardized test scores), and was unable to establish causal predominance between AA and SC. On the basis of findings from these three studies, Marsh (1990c) argued for support of his previous contention (1987a) that prior academic SC is more likely to affect subsequent achievement if achievement is inferred from school grades, rather than from standardized test scores.

To address these issues, Marsh (1990c) conducted a four-wave panel study of the Youth in Transition data (N=1,456; Bachman & O'Malley, 1975). Analyses focused on data at Times 1 (early grade 10, 2 (late grade 11), 3 (late grade 12), and 4 (one year after normal high school graduation); three latent constructs were of interest: (a) academic ability (T1 only), as measured by four standardized test scores, (b) academic SC (T1, T2, T4) as measured by responses to multiple self-rating items, and (c) school grades (T1, T2, T3).

Results showed that, although grade averages in Grades 11 and 12 were significantly affected by academic SC measured in the previous year, prior reported grades had no effect on subsequent measurements of academic SC. On the basis that this study represented the most rigorous design to date for testing causal relations among academic SC and AA, Marsh contended that results demonstrated, convincingly, that academic SC is causally predominant over AA. Finally, in a more recent study of two cohorts (grade 3/4; grade 6/7) of Norwegian children, Skaalvik and Hagtvet (1990) determined causal predominance of AA over SC for the younger children, albeit no such evidence for the older children.

The present study represents a further attempt to determine if, in fact, it is possible to establish causal direction between academic SC and AA. More specifically, the purpose of the study was to determine whether academic SC, and the subject-matter SCs of english and math serve to generate academic, english, and math achievements, respectively. This study of causal predominance was tested at each of three grade levels (3, 7, 11). To the best of my knowledge, this is the first longitudinal SEM study of causal relations to incorporate, into a single investigation, multiple measures of the achievement scores for each of three different age groups.

Method

Sample and Procedures

At two time-points within the same school year (October, May), data were collected from seven elementary and two secondary schools located in suburban areas of a large central Canadian city. Of approximately 400 students sampled at each of three grade levels, parental consent was received for 252 grade 3 (mean age=8 years), 290 grade 7 (mean age=12 years), and 335 grade 11 (mean age=16 years) students, respectively. By necessity, selection of subjects was based on classroom units; only children for whom signed consent was received participated in the study.

Because we considered it important that self-perceptions of math

competence reflected present rather than past recollections of math achievement, data collection at the high school level targeted math classes; subject content was consistent across classes and schools. On the other hand, because english represented a compulsory subject area, the same type of selection process was unnecessary; again, the curriculum was common to all classes.

For each class, data relevant to academic SC and AA measures were collected by trained research assistants during one class period. At each grade level, completion of the SC self-report instrument was preceded by a thorough explanation of the response format, accompanied by a demonstration of several examples. For purposes of additional clarification at the grade 3 level, items were read aloud. Teachers were given a three-day period in which to complete a rating scale for each eligible student in their class.

It is important to note that data for the present study were drawn from a larger research project designed to test the structure of social SC (Byrne & Shavelson, in press). Because (a) the design of the larger study demanded the administration of numerous and structurally diverse measures, and (b) school officials imposed exceedingly rigid time constraints on the data collection process, it was necessary to tailor many of the instruments accordingly. Where applicable, these considerations are noted below.

Measuring Instruments

The three facets of SC relevant to the study (academic, english, math) were measured using appropriate subscales from the Self Description Questionnaires (SDQs) I, II, and III (Marsh, 1992a, 1992b, 1992c). The SDQ-I is appropriate for use with preadolescents (grades 2-6), and uses a 5-point Likert-type scale format that ranges from "false" to "true," indicating the extent to which respondents agree or disagree with self-descriptive statements related to their academic competence. The SDQ-II is designed for use with early adolescents (grades 7-11) and incorporates a 6-point Likert scaling format (false - true); the SDQ-III is used with late adolescents (grade 11-

college), and is based on an 8-point Likert scaling format (definitely false - definitely true). Subsequent to their development in the early 1980's, findings from a plethora of construct validity studies have provided ample evidence supporting the psychometric soundness of all three SDQ instruments (for reviews, see Byrne, 1996; Marsh, 1992a, 1992b, 1992c). Due to data collection time constraints imposed by school authorities, only four items from each selected SDQ subscale were used; selection targeted those items having the highest reported reliability values (see Marsh, 1992a, 1992b, 1992c).

Addressing the limitation of grades as the sole measure of AA noted with respect to several studies (e.g., Byrne & Shavelson, 1986; Shavelson & Bolus, 1982), the present study used multiple measures of academic performance in general, and as it related to the subject areas of english and math. As might be expected, these measures necessarily varied somewhat for each grade level; the measures were as follows:

Standardized achievement tests. The Multilevel Academic Survey Test (MAST: Howell, Zucker, & Morehead, 1985), a standardized measure of overall reading and math, was used to measure AA in these subject areas for children in grades 3 and 7. The Primary Form, designed for use with kindergarten through grade 3 children, was used for the preadolescents; the Short Form, appropriate for children in the higher grades, was used with the early adolescents. The Reading component of both MAST forms requires students to read a short passage after which they respond to multiple-choice (3-alternative) items that tap their comprehension of its content. The math component requires respondents to complete a set of numerical operations which are later scored by the administrator; whereas these operations were based on only whole numbers for grade 3 children, they included both whole and fractional numbers for grade 7 children.

With respect to the Reading Comprehension component of the MAST, Howell et al. (1985) have reported predictive validity coefficients of .74 for both grades 3 and 7 based on related scores from the Iowa Test of Basic Skills, as

well as test-retest (2-week lag; r=.72 grade 3; r=.75 grade 7) and KR-20 internal consistency (r=.90 grade 3; r=.92) reliabilities. Psychometric properties relative to the Math component have revealed predictive validity coefficients of .68 (grade 3) and .72 (grade 7), test-retest reliabilities of .64 (grade 3) and .72 (grade 7) over a 2-week lag, and KR-20 internal consistency reliabilities of .59 and .66 for grades 3 and 7, respectively (Howell et al.).

The Canadian Adult Achievement Test (Form C) (CAAT; Canada Employment and Immigration Commission, 1988), an adaptation of the Adult Basic Learning Examination (ABLE; Karlsen & Gardner, 1986), was used as the standardized measure for grade 11 students. Form C of the CAAT is designed for persons having completed at least eight years of formal education, but who have not yet graduated from high school.

Addressing the focus of the present study, and paralleling achievement measurement for elementary school children, only the Reading Comprehension and Math subscales were used from a total of eight available. The Reading Comprehension component required students to first read a passage of material and then respond to a series of multiple-choice (4-alternative) items bearing on its content. Similarly, the Math component also employed a multiple-choice (5-alternative) format. Following the completion of each numerical operation, respondents select the alternative they consider to be the correct answer.

Unfortunately, information bearing on the psychometric properties of this instrument are meagre. The authors of the CAAT contend that it constitutes a representative sample of the skills, knowledge, and comprehension consistent with the level of mastery expected for the intended population and, as such, demonstrates sound content validity. They also report KR-20 internal consistency reliability coefficients of .91 and .92 for the Reading Comprehension and Math subscales, respectively.

Grades. The original intent of the study was to collect school grades as one measure of achievement for both english and math at each of the three

grade levels. Unfortunately, immediately prior to the collection of data, government legislation was introduced that banned public access to information related to school records. Whereas this new legislation precluded collection of student grades for children in grades 3 and 7, we were nonetheless successful in obtaining this information at the high school level. As such, grades for these students represented their final marks in english and math for both the Fall and Spring terms.

Self-ratings. To compensate for the inability to obtain grade data for elementary school children, separate self-ratings of achievement in english and math were collected. However, given (a) the low saliency of academic success and failure for early preadolescent children (i.e., grade 3, Harter, 1992), (b) the tendency for preadolescents to inflate their sense of adequacy (Harter, 1988), and (c) the fact that for the present sample of grade 3 children, subject-specific achievement evaluations were recorded as verbal descriptions, rather than as letter or numerical grades, self-ratings were limited to only children in grades 7 and 11.

Presented with a scale ranging from 1 to 10, students were asked to indicate the grade-point value that best summarized their achievement in english and in math. To assist them in this evaluation, each grade-point value was accompanied by a letter grade, a numerical grade, and a verbal description. For example, a grade-point value of 10 was equated with a letter grade of A+, a numerical grade of 90-100, and a verbal description of "exceptional"; a grade-point value of 3 was equated with a letter grade of D+, a numerical grade of 56-59, and a verbal description of "adequate".

Teacher ratings. Ratings of general academic performance was based on Harter's Teacher Rating Scale; ratings of performance in english and math was based on the same form as that described above for students. Only one minor modification to the instructions was made to render them appropriate for responses by teachers. That is to say, the teacher was instructed to indicate the grade-point value which best summarized "this student's" achievement in english and math.

Analysis of the Data

Structural equation modeling (SEM) procedures were used to test hypotheses related to causal direction. All procedures were based on the analysis of covariance structures using the EQS program (Bentler, 1995). Analyses were conducted in three stages for each grade separately. First, to facilitate interpretation, negative items were reversed such that high scores represented highly positive perceptions. Second, given the known possibility of memory effects associated with the measures across time, a four-factor model (SC, AA at two time points) was tested for goodness-of-fit at each grade level for general academic SC/achievement relations, and each of the two subject-specific SC/achievement relations. Finally, based on the best-fitting model from these analyses, causal predominance between the SC and achievement constructs was tested by comparing the fit of a model in which the two competing structural paths (SC -> AA; AA ->SC) were specified as free, with one in which these paths were constrained to be equal. (For greater elaboration of both the conceptual underpinning and application of these procedures, see Byrne, 1994a, 1998.) A more detailed description of both the hypothesized models and goodness-of-fit criteria used in the present study now follows.

Hypothesized models. For each set of SC/AA relations at each grade level, the SEM model hypothesized the following conditions a priori: (a) two competing structural regression paths across time; one postulating that AA causes SC (F2 -> F3), and the other, that SC causes AA (F1 -> F4). These models are presented in Figures 1, 2, and 3 for grades 3, 7, and 11, respectively; (b) each indicator variable would have a non-zero loading on the SC or AA factor it was designed to measure, and zero loadings on all other factors; (c) the SC and AA factors, at Time 1 only, would be correlated;[1] and (d) measurement error terms would be uncorrelated, both within and across time points.

Insert Figures 1, 2, and 3 about here

Goodness-of-fit criteria. Evaluation of model fit was based on multiple criteria that took substantive, statistical, and practical fit into account. Specifically, these criteria included: (a) the substantive meaningfulness of the model (MacCallum, 1986), (b) the χ^2 likelihood ratio statistic, (c) the Satorra-Bentler corrected χ^2 statistic (S-B χ^2 ; Satorra & Bentler, 1988, and (d) the Robust Comparative Fit Index, based on the S-B χ^2 statistic (CFI; Bentler, 1990, 1995).

The χ^2 likelihood ratio statistic, in practice, is more useful when regarded as a **measure** of fit, rather than as a **test** statistic (Jqreskog & Sorbom, 1993). As such, the χ^2 value measures the closeness of fit between the sample covariance matrix and the fitted covariance matrix, serving therefore as an indicator of overall model fit. However, given the known dependency of the χ^2 statistic on sample size and the grounding of covariance structure analysis in large sample theory, findings typically indicate a need to modify the model in order to better fit the data (Joreskog & Sorbom, 1993). As a consequence, it has become customary to base evaluation of a model on practical indices of fit such as the CFI described below.

The S-B χ^2 serves as a correction for the χ^2 statistic when distributional assumptions are violated. Its computation takes into account the model, the estimation method, and the sample kurtosis values. The S-B χ^2 has been shown to be the most reliable test statistic for evaluating covariance structure models under various distributions and sample sizes (Hu, Bentler, & Kano, 1992).

The CFI, a revised version of the Bentler-Bonett (1980) normed fit index that adjusts for degrees of freedom, ranges in value from zero to 1.00. It is derived from the comparison of a restricted model (i.e., one in which

structure is imposed on the data) with an independence (or null) model (one in which all correlations among variables are zero) in the determination of goodness-of-fit. A CFI value of .90 has served as the rule-of-thumb lower limit cutpoint of acceptable fit. Computation of the Robust CFI (CFI)* is based on the S-B χ^2 values, rather than on the uncorrected χ^2 values. Because evaluation of model fit was based on the S-B χ^2 statistic, the CFI*, rather than the CFI, was used as the index of practical fit.

Given (a) the propensity of pairwise deleted correlational data to yield a non-positive definite covariance matrix (Bentler & Chou, 1987; Kaplan, 1990), and (b) the caveat that current structural modeling methods were designed for use with complete data (Bentler & Chou, 1987), it was considered most appropriate to base analyses on complete data. To this end, all cases missing four or more item scores were deleted from the analyses; for those having less than four missing item scores, mean imputation was invoked. Results yielded final sample sizes of 210, 222, and 232 for grades 3, 7, and 11, respectively.

Results

Preliminary analyses identified three multivariate outliers in both the grades 7 and 11 samples; these cases were subsequently deleted from all analyses thereby resulting in final sample sizes of 219 and 229 for grades 7 and 11, respectively. Study findings are now presented separately by grade level.

Grade 3

Test of Model Fit

As noted earlier, one expected finding related to this study, was the likelihood of error covariances (commonly referred to as correlated errors) across time for particular indicator variables. Furthermore, given that correlated errors among items of the same measuring instrument are not uncommon in psychological research, these error parameters also were not unexpected. Indeed, both types of correlated errors contributed to the malfit

of initially hypothesized models. In determining the best-fitting model of SC/achievement relations for grade 3 (and for grades 7 and 11), then, malfitting initial models were respecified to include error covariances as identified by the LMTest, but only if they were deemed to be substantively meaningful parameters. A summary of the model-fitting statistics and incorporated error covariances for Grade 3 are presented in Table 1.

Insert Table 1 about here

Academic SC/AA. In reviewing results bearing on these relations, it is evident that the initially hypothesized model represented a very poor fit to the data (CFI^{*}=.74). In EQS, such misspecification can be determined from the Lagrange Multiplier (LM) multivariate χ^2 statistics which identify fixed parameters that, if freely estimated, would lead to a better-fitting model. It is important to note, however, that while these statistically-driven modification indices are helpful in targeting malfitting parameters in the model, any respecification should only be made in light of sound substantive rationale. In the present case, the misspecified parameters represented correlated errors across time for both teacher rating scores, and one correlated error between teacher ratings of AA at Time 1 only (see footnote a, Table 1). Although the first two parameters reasonably derive from memory effects, the latter likely represents redundant item content. Indeed such parameters often occur within and across subscales of the same measuring instrument, and are commonly found with respect to attitude scales in general (see e.g., Byrne, 1993, 1994b; Newcomb, Huba, & Bentler, 1986) and SC measures in particular (see e.g., Byrne & Shavelson, 1986, 1987, 1996; Byrne & Worth Gavin, 1996). On the basis of both distinctively large LMχ^2 statistics and a valid psychometric rationale then, the initial academic SC/AA model for grade 3 children was respecified to include the error correlations noted above. As can be seen in Table 1, this nested model (Model 2) yielded a statistically

significant difference in fit from that of the initially hypothesized model (Model 1) ($\Delta\chi^2_{(3)}$=128.21).[2] Admittedly, the fit of Model 2 is marginally adequate. However, although the LMTest statistics identified two additional parameters that if set free would lead to a statistically better-fitting model, these parameters involved further error correlations among the teacher rating scale indicators which, when specified and estimated, led to statistically unreasonable parameter estimates; they were therefore not included in the model.

In testing for predominance of causal direction between academic SC and AA, the final model was specified with both structural paths constrained equal (Model 3). The fit of this model was then compared with Model 2 in which both parameters were freely estimated. As shown in Table 1, this comparison led to a statistically significant difference in model fit ($\Delta\chi^2_{(1)}$=53.56) thereby indicating evidence of inequality between the the two structural paths. A review of the parameter estimates revealed AA to be causally predominant over academic SC (.871 vs .535).

English SC/english achievement. Turning now to models representing english SC/english achievement relations, we find, once again, an ill-fitting hypothesized model (CFI*=.69). Respecification of the model to include (a) error correlations across time for the standardized english test and teacher ratings of english achievement, and (b) an error correlation between items #29 and 40 of the SDQ1 at Time 2 only, led to a statistically better-fitting final model (CFI*=.94) and one that was significantly different from Model 1 ($\Delta\chi^2_{(3)}$=241.44). The test for causal predominance between english SC and english achievement revealed no statistically significant difference between the two structural paths ($\Delta\chi^2_{(1)}$=0.91).

Math SC/mathematics achievement. Finally, as shown in Table 1, two error correlations were additionally specified for the initial model (see footnote b), thereby leading to a final better-fitting model of math SC/math achievement relations (CFI*=.94; $\Delta\chi^2_{(2)}$=161.60). The test for causal predominance yielded a

statistically significant difference between Models 2 and 3 ($\Delta\chi^2(1)=14.53$), with causal direction shown to flow from math achievement to math SC (.698 vs .562).

Insert Table 2 about here

Grade 7

Academic SC/AA. As shown in Table 2, establishment of a well-fitting model of academic SC/AA relations for grade 7 children involved the additional incorporation of three error correlations into the initially hypothesized model (see footnote a). The test for causal predominance between academic SC and AA revealed no significant differences between Models 2 and 3 ($\Delta\chi^2_{(1)}=0.24$) thereby indicating that such predominance could not be established.

English SC/english achievement. Five error correlations were involved in the determination of a well-fitting model of english SC/achievement relations ($CFI^*=.93$); three of these parameters represented error correlations between SDQ2 items (#10, #34, #46), the remaining two were associated with standardized english test scores and teacher ratings across time. In contrast to the grade 3 findings, causal predominance could be established ($\Delta\chi^2_{(1)}=17.50$), with direction of cause flowing from english SC to english achievement (.885 vs .815).

Math SC/math achievement. The final best-fitting model of math SC/achievement relations specified two error correlations; these were associated with teacher ratings of math and the SDQ2 item #43 across time. As was the case for academic SC/AA relations, direction of cause between math SC and math achievement could not be established ($\Delta\chi^2_{(1)}=0.10$).

Insert Table 3 about here

Grade 11

Academic SC/AA. In reviewing the results presented in Table 3, we can see that five error correlations were added to the specification of the initally hypothesized model in order to edetermine the better-fitting final model of academic SC/AA relations (CFI^{*}=.91; for details, see footnote a). Evidence of causal predominance was established ($\Delta\chi^2_{(1)}$=15.96), with results showing direction of flow from AA to academic SC (.881 vs .693).

English SC/english achievement. Determination of a final best-fitting model of english SC/achievement relations (CFI^{*}=.95) involved the specification of four error correlations; all were associated with SDQ3 items across time (#10, #22, #34, #43). Direction of causal flow was firmly established with predominance being attributed to the english SC -> english achievement path (.821 vs .608).

Math SC/math achievement. Finally, from Table 3, we see that five error correlations were specified in order to attain a well-fitting final model (CFI^{*}=.93; for details, see footnote c). Causal predominance, once again, was clearly established ($\Delta\chi^2_{(1)}$=18.67) with direction of cause flowing from math achievement to math SC (.807 vs .756).

A summary of results related to causal predominance between each self-concept dimension and its related achievement construct is presented in Table 4.

Insert Table 4 about here

Discussion

The present study attempted to establish direction of causal influence between academic SC and AA in general, as well as between the more specific dimensions of english and math SCs, and their related achievement scores, for children in grades 3, 7, and 11. Of the nine sets of parameters tested for causal predominance (three at each grade level), direction of cause was

established for six; inconclusive findings resulted for english SC/english achievement relations for grade 3 children, and for academic SC/AA and math SC/math achievement relations for grade 7 children. For both grades 3 and 11, tests of causal predominance involving academic SC/achievement relations, and math SC/achievement relations demonstrated a clear flow of causality **from** achievement **to** self-concept. In contrast, SC/achievement relations associated with the subject area of english demonstrated the reverse pattern (i.e., **from** english SC **to** english achievement).

Presented with these findings, three primary questions cry out for answers: (a) why is it that causal predominance between SC and academic performance can be established for some academic subject areas and not for others?; (b) why is it that causal predominance can be established for some grade levels and not for others?; and (c) why is it that, when causal predominance can be established, the direction of cause flows **from** achievement **to** SC for academic SC/AA and math SC/achievement relations, but, flows **from** SC **to** achievement for english SC/achievement relations?

Answers to the first question are likely linked to the issue of age. For example, Harter (1988, 1990) has noted that, across the childhood years, children's self views change from a focus on concrete, observable aspects of the self, to psychological traits, and finally to abstractions and psychological processes; these changes derive, in part, from concomitant changes in their cognitive processing skills. Self-concept research for young children has shown them to be able to make reliable judgements of their academic competence, but only if these domains are depicted as concrete, observable behaviors (Harter, 1988, 1990). That causal predominance, for grade 3 children, could be established for academic SC/AA and math SC/achievement relations likely stems from the fact that academic work in general, and math work in particular, are relatively well-defined and concrete phenomena. As a consequence, children are able to formulate self-perceptions of competency related to these areas. On the other hand, given the broad spectrum of course

content possible with respect to english, it seems reasonable to assume that this academic subject may be somewhat less well-defined and, hence, a less concrete phenomenon on which to base their self-perceptions.

In contrast to the grade 3 results, however, those for grades 7 and 11 children yielded evidence of causal predominance with respect to english SC/achievement relations. This finding is likely tied to the more sophisticated cognitive processing skills of the older children. As such, it seems reasonable that these early and late adolescents would have less difficulty comprehending the parameters of english as a subject area than would preadolescents. This being the case, it follows that they would have less difficulty in formulating self-perceptions of their competence in this academic area.

The question of why causal predominance can be established for some grade levels and not for others is likely answered, once again, by the age factor. Consistent with previous research (e.g., Newman, 1984; Skaalvik & Hagtvet, 1990; Wigfield & Karpathian, 1991), the present study determined grade 7 children to march, apparently, to a different drummer than children in grades 3 and 11. Whereas both Newman (1984), and Skaalvik and Hagtvet (1990) were able to establish evidence of causal predominance between academic SCs and their matching achievement scores for younger children (grades 2-5), such evidence was not forthcoming for older children (grades 6-10). Relatedly, several studies of mean group differences have shown lower SCs for children in middle childhood than for young children and late adolescents (e.g., Eccles, Wigfield, Harold, & Blumenfield, 1991; Marsh, 1989). These findings suggest that SC tends to be less stable for children in the middle years than it is for children in the later stages of childhood. This fact, then, may partially explain why results bearing on causal predominance were completely reversed for children in grades 3 and 7. Indeed, this factor of instability prompted Wigfield and Karpathian (1991) to conclude that, in middle childhood, SC/AA relations are likely reciprocal.

Finally, the question of why direction of cause flows **from** achievement **to** SC for academic SC/AA and math SC/achievement relations, but **from** SC **to** achievement for english SC/achievement relations, is a conundrum and not easily answered. One possibility may be that academic work in general, and math work in particular, are clearly defined constructs and therefore clearly conceptualized by students. As a consequence, evaluation of self-performance in these areas can be meaningfully interpreted and therefore likely forms the basis of one's perception of self (i.e., achievement determines self-perception relative to the area of achievement). In contrast, english as a subject, can vary widely from school to school, teacher to teacher, and grade to grade. Because its conceptualization is more nebulous than the other two areas, it does not serve as a basis upon which to formulate one's perception of self. Rather, the causal direction flows in the reverse order. However, these comments are purely speculative; more further investigative work is needed in order to untangle this intriguing inconsistency.

Findings from this study have added another piece to the puzzle concerned with direction of cause between academic SC facets and their matching achievement scores. The fact that causal predominance related to english SC/achievement relations differed from that for scholastic achievement in general, and math achievement in particular, suggests that direction of cause may be a function of subject area, rather than, or in addition to age. Further replication work that involves longitudinal studies based on sample sizes larger than those of the present study are needed in order to untangle the intriguing and complex web of factors bearing on causal relations between SC and AA.

References

Bachman, J.G. & O'Malley, P.M. (1975). Youth in transition, data file documentation (Vol 2). Ann Arbor MI: Institute for Social Research.

Bentler, P.M. (1990). Comparative fit indices in structural models. Psychological Bulletin, 107, 238-246.

Bentler, P.M. (1995). EQS: Structural equations program manual. Los Angeles, CA: Multivariate Software Inc.

Bentler, P.M. & Bonett, D.G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. Psychological Bulletin, 88, 588-606.

Bentler, P.M. & Chou, C.P. (1987). Practical issues in structural modeling. Sociological Methods & Research, 16, 78-117.

Byrne, B.M. (1984). The general/academic self-concept nomological network: A review of construct validation research. Review of Educational Research, 54, 427-456.

Byrne, B.M. (1986). Self-concept/academic achievement relations: An investigation of dimensionality, stability, and causality. Canadian Journal of Behavioural Science, 18, 173-186.

Byrne, B.M. (1993). The Maslach Burnout Inventory: Testing for factorial validity and invariance across elementary, intermediate, and secondary teachers. Journal of Occupational and Organizational Psychology, 66, 197-212.

Byrne, B.M. (1994a). Structural equation modeling with EQS and EQS/Windows: Basic concepts, applications, and programming. Newbury Park, Sage.

Byrne, B.M. (1994b). Testing for the factorial validity, replication, and invariance of a measuring instrument: A paradigmatic application based on the Maslach Burnout Inventory. Multivariate Behavioral Research, 29, 289-311.

Byrne, B.M. (1996). Measuring self-concept across the lifespan: Issues and Instrumentation. Washington DC: American Psychological Association.

Byrne, B.M. (1998). Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming. Mahwah NJ: Erlbaum.

Byrne, B.M. & Shavelson, R.J. (1986). On the structure of adolescent self-concept. Journal of Educational Psychology, 78, 474-481.

Byrne, B.M. & Shavelson, R.J. (1996). On the structure of social self-concept for pre-, early, and late adolescents. Journal of Personality and Social Psychology, 70, 599-613.

Byrne, B.M. & Worth Gavin, D.A. (1996). The Shavelson model revisited: Testing for the structure of academic self-concept across pre-, early, and late adolescents. Journal of Educational Psychology, 88, 215-228.

Calsyn, R.J. & Kenny, D.A. (1977). Self-concept of ability and perceived evaluation of others: Cause or effect of academic achievement? Journal of Educational Psychology, 69, 136-145.

Canada Employment and Immigration Commission (1988). CAAT: Canadian Adult Achievement Test. Toronto: Harcourt Brace Jovanovich.

Chapman, J.W., Cullen, J.L., Boersma, F.J., & Maguire, T.D. (1981). Affective variables and school achievement: A study of possible causal influences. Canadian Journal of Behavioural Science, 13, 181-192.

Cook, T.D. & Campbell, D.T. (1979). Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally.

Eccles, J.S., Wigfield, A., Harold, R., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. Child Development, 64, 830-847.

Harter, S. (1985). Manual for the Self-Perception Profile for Children. Denver, CO: University of Denver.

Harter, S. (1988). Developmental processes in the construction of the self. In J. T.D. Yawkey & J.E. Johnson (Eds.), Integrative processes and socialization: Early to middle childhood (pp. 45-78). Hillsdale NJ: Erlbaum.

Harter, S. (1990). Issues in the assessment of the self-concept of children and adolescents. In A.M. LaGreca (Ed.), Through the eyes of the child: Obtaining self-reports from children and adolescents (pp. 292-325). Boston: Allyn and Bacon.

Hattie, J. (1992). Self-concept. Hillsdale N.J.: Erlbaum.

Howell, K.W., Zucker, S.H., & Morehead, M.K. (1985). MAST: Multilevel Academic Survey Test. Columbus OH: Merrill.

Hu, L., Bentler, P.M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? Psychological Bulletin, 112, 351-362.

Joreskog, K.G. & Sorbom D. (1993). LISREL 8: Structural equation modeling with the SIMPLIS command language. Hillsdale, N.J.: Erlbaum.

Kaplan, D. (1990). Evaluating and modifying covariance structure models: A review and recommendation. Multivariate Behavioral Research, 25, 137- 155.

Karlsen, B & Gardner, E.F. (1986). ABLE: The Adult Basic Learning Examination. San Antonio TX: The Psychological Corporation.

Kenny, D.A. (1979). Correlation and causality. New York: Wiley.

MacCallum, R. (1986). Specification searches in covariance structure modeling. Psychological Bulletin, 100, 107-120.

Marsh, H.W. (1987). The big-fish-little-pond effect on academic self-concept. Journal of Educational Psychology, 79, 280-295.

Marsh, H.W. (1988). Causal effects of academic self-concept on academic achievement: A reanalysis of Newman (1984). Journal of Experimental Education, 56, 100-104.

Marsh, H.W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to to early childhood. Journal of Educational Psychology, 81, 417-430.

Marsh, H.W. (1990). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. Journal of Educational Psychology, 82, 646-656.

Marsh, H.W. (1992a). Self Description Questionnaire (SDQ) I: A theoretical and empirical basis for the measurement of multiple dimensions of preadolescent self-concept: A test manual and research monograph. Macarthur, NSW Australia: Faculty of Education, University of Western Sydney.

Marsh, H.W. (1992b). Self Description Questionnaire (SDQ) II: A theoretical and empirical basis for the measurement of multiple dimensions of adolescent self-concept: An interim test manual and research monograph. Macarthur, NSW Australia: Faculty of Education, University of Western Sydney.

Marsh, H.W. (1992c). Self Description Questionnaire (SDQ) III: A theoretical and empirical basis for the measurement of multiple dimensions of late adolescent self-concept: An interim test manual and research monograph. Macarthur, NSW Australia: Faculty of Education, University of Western Sydney.

Moyer, D.C. (1980). Academic achievement, self-concept and locus of control: A causal analysis of the National Longitudinal Study. Dissertation Abstracts International, 40, 4496A. (University Microfilms No. 8003756).

Newcomb, M.D., Huba, G.T., & Bentler, P.M. (1986). Determinants of sexual and dating behaviors among adolescents. Journal of Personality and Social Psychology, 50, 428-438.

Newman, R.S. (1984). Children's achievement and self-evaluations in mathematics: A longitudinal study. Journal of Educational Psychology, 76, 857-873.

Pottebaum, K.A., Keith, T.Z., & Ehly, S.W. (1986). Is there a causal relation between self-concept and academic achievement? Journal of Educational Research, 79, 140-144.

Rogosa, D. (1980). A critique of cross-lagged correlation. Psychological Bulletin, 88, 245-258.

Satorra, A. & Bentler, P.M. (1988). Scaling corrections for chi-square statistics in covariance structure analysis. American Statistical Association 1988 Proceedings of the Business and Economic Sections (pp. 308-313). Alexandria VA: American Statistical Association.

Scheirer, M.A. & Kraut, R.E. (1979). Increasing educational achievement via self-concept change. Review of Educational Research, 49, 131-150.

Shavelson, R.J. & Bolus, R. (1982). Self-concept: The interplay of theory and methods. Journal of Educational Psychology, 74, 3-17.

Skaalvik, E.M. & Hagtvet, K.A. (1990). Academic achievement and self-concept: An analysis of causal predominance in a developmental perspective. Journal of Personality and Social Psychology, 58, 292-307.

Watkins, D. & Astilla, E. (1987). Causal dominance among self-concept, locus of control, and academic achievement. The Journal of Psychology, 120, 627-633.

West, C.K., Fish, J.A., & Stevens, R.J. (1980). General self-concept, self-concept of academic ability and school achievement: Implications for "causes" of self-concept. The Australian Journal of Education, 24, 194-213.

Wigfield, A. & Karpathian, M. (1991). Who am I and what can I do? Children's self-concepts and motivation in achievement situations. Educational Psychologist, 26, 233-261.

Footnotes

1. Because correlations at Time 2 must remain to be explained by the data, it is incorrect to include the specification of variances and covariances for dependent variables in a SEM model. Arrows representative of these parameters have been excluded from Figures 1-3 for purposes of clarity.

2. The difference between χ^2 values of nested models is distributed as a χ^2 statistic, with degrees of freedom equal to the difference in degrees of freedom related to each of the models being compared. This value ($\Delta\chi^2_{(df)}$) therefore serves as a means to assessing the viability of a respecified model.

Table 1

Summary Statistics Related to Model-fitting and Tests of Equality for Grade 3

Academic Self-concept vs Academic Achievemnt

Model	χ^2	df	S-Bχ^2	CFI*	Difference in ΔS-Bχ^2	Δdf
1 Initial	340.32	51	251.97	.74		
2 Final (3 error correlations)[a]	152.07	48	123.76	.90	128.56***	3
3 Model 2 with structural paths constrained equal	215.80	49	177.32	.83		
4 Comparison of Models 2 and 3					53.56***	1

English Self-concept vs English Achievement

1 Initial	405.83	51	351.39	.69		
2 Final (3 error correlations)[b]	121.14	48	109.95	.94	241.44***	1
3 Model 2 with structural paths constrained equal	121.74	49	110.86	.94		
4 Comparison of Models 2 and 3					0.91	1

Math Self-concept vs Math Achievement

1 Initial	313.50	51	277.63	.78		
2 Final (2 error correlations)[c]	130.56	49	116.03	.94	161.60***	2
3 Model 2 with structural paths constrained equal	135.98	50	130.56	.93		
4 Comparison of Models 2 and 3					14.53***	1

Table 1 Cont'd

*** $p<.001$

S-Bχ^2=Satorra-Bentler Scaled Statistic; CFI*=Scaled Comparison Fit Index; AA=Academic Achievement; ESC=english self-concept

[a] teacher ratings of AA across Times 1 and 2 (both items); teacher ratings of AA at Time 1

[b] standardized english test scores across Times 1 and 2; teacher ratings of english achievement across Times 1 and 2; correlation between SDQ1 ESC items #29 and 40 at Time 2

[c] standardized math test scores across Times 1 and 2; teacher ratings of math across Times 1 and 2

Table 2

Summary Statistics Related to Model-fitting and Tests of Equality for Grade 7

Model	χ^2	df	S-Bχ^2	CFI*	Difference in ΔS-Bχ^2	Δdf
Academic Self-concept vs Academic Achievemnt						
1 Initial	300.78	51	241.64	.81		
2 Final (3 error correlations)[a]	152.03	48	123.09	.93	118.55***	3
3 Model 2 with structural paths constrained equal	152.33	49	122.85	.93		
4 Comparison of Models 2 and 3					0.24	1
English Self-concept vs English Achievement						
1 Initial	314.77	74	245.12	.73		
2 Final (5 error correlations)[b]	141.87	69	112.39	.93	132.73***	5
3 Model 2 with structural paths constrained equal	163.64	70	129.89	.91		
4 Comparison of Models 2 and 3					17.50***	1
Math Self-concept vs Math Achievement						
1 Initial	292.11	74	270.22	.90		
2 Final (2 error correlations)[c]	265.82	72	223.11	.93	47.11***	2
3 Model 2 with structural paths constrained equal	265.82	73	223.21	.93		
4 Comparison of Models 2 and 3					0.10	1

Table 2 Cont'd

*** $p<.001$

S-Bχ^2=Satorra-Bentler Scaled Statistic; CFI*=Scaled Comparison Fit Index; AA=Academic Achievement; ASC=academic self-concept; ESC=english self-concept; MSC=math self-concept

[a] teacher ratings of AA at Time 1; SDQ2 ASC items #4 and #16 across Times 1 and 2

[b] standardized english test scores across Times 1 and 2; teacher ratings of english achievement across Times 1 and 2; SDQ2 ESC items #10, #34, and #46 across Times 1 and 2

[c] teacher ratings of math across Times 1 and 2; SDQ2 MSC item #43 across Times 1 and 2

Table 3

Summary Statistics Related to Model-fitting and Tests of Equality for Grade 11

Academic Self-concept vs Academic Achievemnt

Model	χ^2	df	S-Bχ^2	CFI*	Difference in ΔS-Bχ^2	Δdf
1 Initial	409.65	51	312.55	.76		
2 Final (5 error correlations)[a]	183.52	46	145.58	.91	166.97***	5
3 Model 2 with structural paths constrained equal	209.69	47	161.54	.89		
4 Comparison of Models 2 and 3					15.96***	1

English Self-concept vs English Achievement

1 Initial	746.40	101	659.82	.80		
2 Final (4 error correlations)[b]	283.01	97	245.35	.95	414.47***	4
3 Model 2 with structural paths constrained equal	428.88	98	368.98	.90		
4 Comparison of Models 2 and 3					123.63***	1

Math Self-concept vs Math Achievement

1 Initial	703.78	101	615.28	.85		
2 Final (5 error correlations)[c]	408.32	96	344.47	.93	270.81***	5
3 Model 2 with structural paths constrained equal	433.87	97	363.14	.92		
4 Comparison of Models 2 and 3					18.67***	1

Table 3 Cont'd

*** $p<.001$

S-Bχ^2=Satorra-Bentler Scaled Statistic; CFI*=Scaled Comparison Fit Index; AA=Academic Achievement; ASC=academic self-concept; ESC=english self-concept; MSC=math self-concept

[a] teacher ratings of AA at Time 1; teacher ratings of AA (1 item) across Times 1 and 2; correlation between teacher rating item # 6 (Time 1) with item 1 (Time 2); SDQ3 ASC items #4 and #28 across Times 1 and 2

[b] SDQ3 ESC items #10, #22, #34, and #43 across Times 1 and 2

[c] standardized math test scores across Times 1 and 2; correlation between math marks and teacher ratings of math at both Time 1 and Time 2; SDQ3 MSC items #7 and #31 across Times 1 and 2

Table 4

Parameter Estimates for Structural Paths

Grade	Structural Path	Estimate	Direction of Cause
3	AA -> ASC	.871	AA -> ASC
	ASC -> AA	.535	
	ENG -> ESC	.710	Inconclusive
	ESC -> ENG	.568	
	MATH -> MSC	.698	MATH -> MSC
	MSC -> MATH	.562	
7	AA -> ASC	.782	Inconclusive
	ASC -> AA	.866	
	ENG -> ESC	.815	ESC -> ENG
	ESC -> ENG	.885	
	MATH -> MSC	.880	Inconclusive
	MSC -> MATH	.954	
11	AA -> ASC	.881	AA -> ASC
	ASC -> AA	.693	
	ENG -> ESC	.608	ESC -> ENG
	ESC -> ENG	.821	
	MATH -> MSC	.807	MATH -> MSC
	MSC -> MATH	.756	

AA = academic achievement; ASC = academic self-concept; ENG = english achievement; ESC = english self-concept; MATH = math achievement; MSC = math self-concept

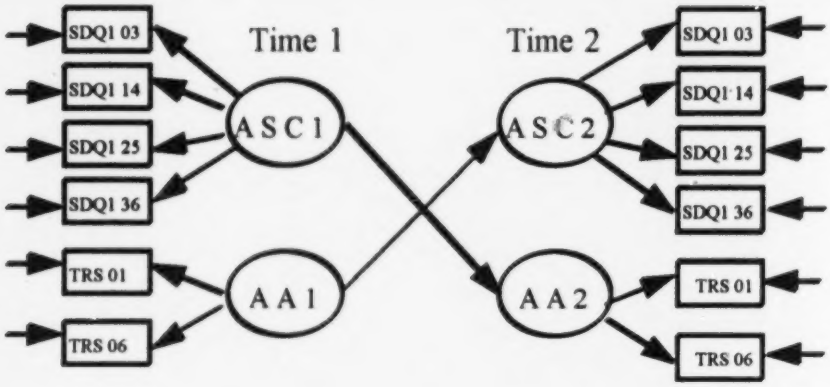
Figure Captions

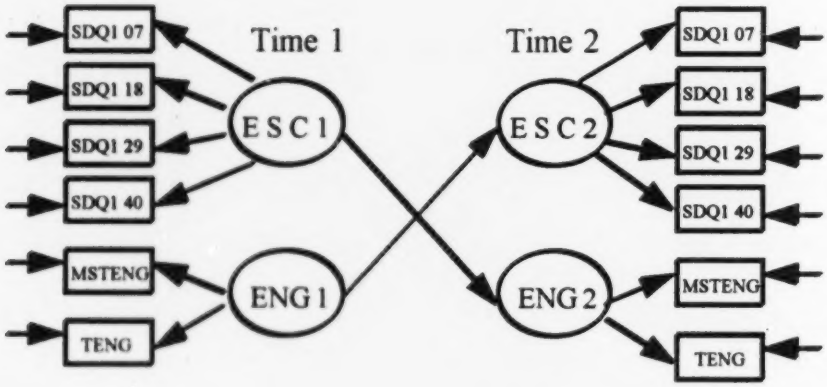
Figure 1. Hypothesized Models of Self-concept/Academic Achievement Relations for Grade 3

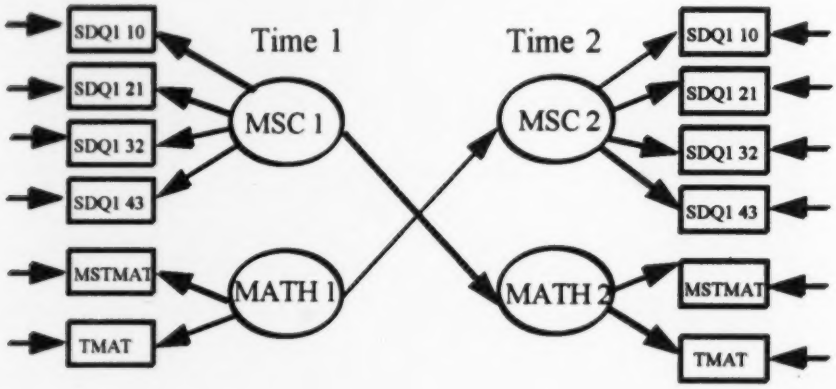
Figure 2. Hypothesized Models of Self-concept/Academic Achievement Relations for Grade 7

Figure 3. Hypothesized Models of Self-concept/Academic Achievement Relations for Grade 7

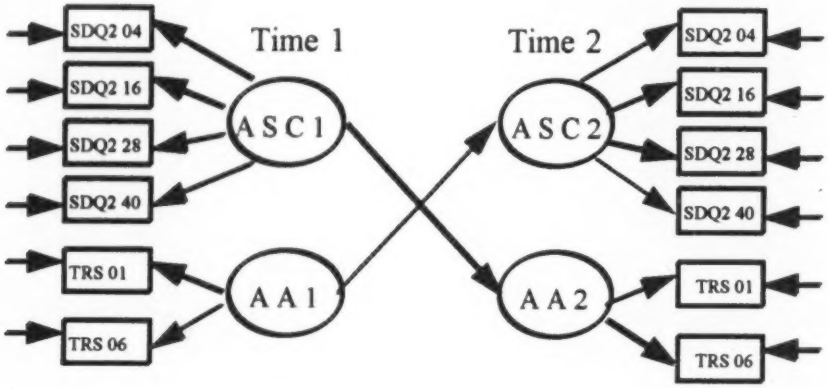
Hypothesized Models - Grade 3

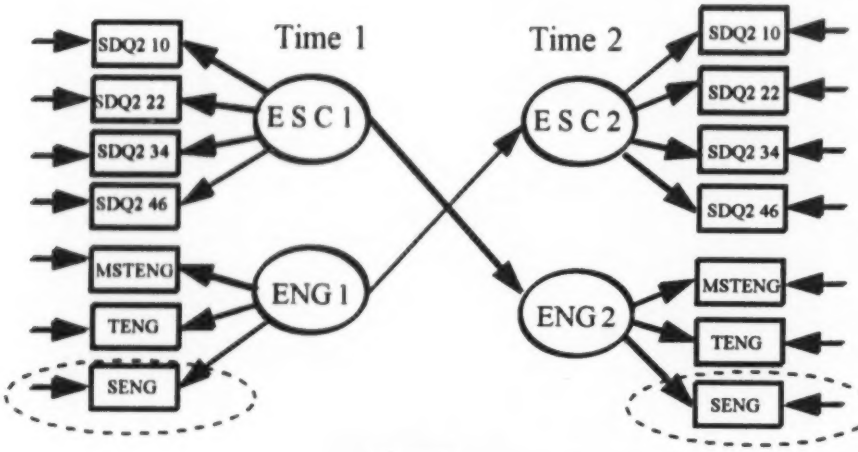


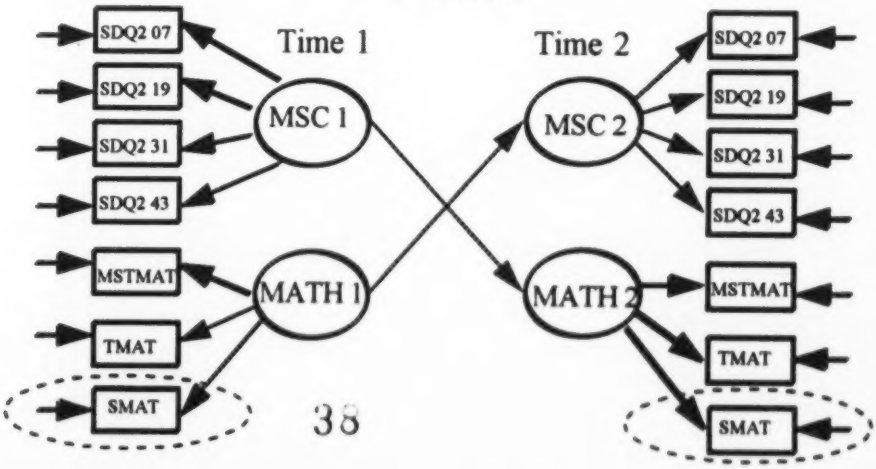




Hypothesized Models - Grade 7







Hypothesized Models - Grade 11